RCA 90,316

CITED BY APPLICANT

Federal Republic of Germany

German Patent and Trademark Office

Patent Application (unexamined) DE 197 41 596 A 1

Int. Cl. [6]:
H 04 R 1/20
H 04 M 1/60
H 04 M 9/08

File Number: 197 41 596.2
Application Date: 20 September 97
Layed Open: 25 March 99

Applicant: Robert Bosch GmbH, 70469 Stuttgart, DE

Inventors: Wietzke, Joachim, Dr., 31141 Hildesheim, DE;
Cornelius, Rainer, 31199 Diekholzen, DE

The following specifications are contained in the documents submitted by the applicant

Process for optimizing the reception of acoustic signals and electric apparatus

A process for optimizing the reception of acoustic signals and an electric apparatus (1), in particular a telecommunication end apparatus, is proposed, which serves for the interference-free reception of acoustic signals. To the electric apparatus (1) at least two microphones (5, 10) are connectable. The microphones (5, 10, 15) transduce received acoustic signals into electric signals. Furthermore, at least one addition element (20, 25) is provided which superimposes the electric signals of connected microphones (5, 10, 15), wherein at least one phase delay element (30, 35, 40, 45) is provided, which delays an electric signal in its phase before the superposition. The phase delay of the at least one phase delay element (30, 35, 40, 45) is selected such that, as a function of the location of the connected microphones (5, 10, 15), for at least one preset site (200), the amplitude of the superposition signal is within a preset range, if at the at least one preset site a sound source (55) outputs acoustic signals.

Specification

Prior Art

The invention builds on a process according to the species of the independent claim 1 and on an electric apparatus according to the species of the independent claim 5.

Electric apparatus which permit voice input are already known in the form of telephone sets. The voice input takes place therein, for example, via a hand set-free speaking microphone.

Advantages of the Invention

The process according to the invention with the characteristics of the independent claim 1 and the electric apparatus according to the invention with the characteristics of the independent claim 5, in comparison, have the advantage that, due to the phase-shifted superposition of the electric microphone output signals, a characteristic directional effect is achieved. In this way, the sensitivity at a preset site in the room is increased such that a sound source disposed at this side can be received especially well by the microphones and an exclusion of noise signal sources at other sites in the room becomes possible. Therefrom results an increased comprehensibility for the human ear in the transmission of voice signals, for example via a telecommunication network, as well as also for a voice processing system in a voice-controlled electric apparatus, such that disturbance effects are from the outset not picked up and, accordingly, also do not need to be suppressed through expensive measures. The word recognition probability of a voice recognition system is correspondingly increased and word analysis simplified. The signal is less falsified by background noises.

Due to the measures listed in the dependent claims, advantageous further

developments and improvements of the process specified in the independent claim 1 and of the electric apparatus specified in the independent claim 5 are possible.

It is of advantage that different phase delays of the at least one phase-delay element are settable. In this way, the setting, independently of location, of a receiving maximum for the superposition signal is possible.

It is especially advantageous that a signal processing unit is provided to which are supplied the electric signals of the microphones and which, as a function of the amplitudes of the electric signals, determines the coordinates of at least one sound source.

Depending on the number and location selection of the microphones, in this way from the received signals two or three-dimensional images of the sound environment can be calculated, such that all sound sources can be determined in terms of location. Based on this information, the phase delay of the at least one phase delay element can subsequently be set such that for a desired sound source a receiving maximum for the superposition signal results.

A further advantage comprises that with the signal processing unit is associated a voice analysis device, and that the voice analysis device carries out a comparison of parameters of the electric signals with voice parameters stored in an associated storage unit and identifies a sound source as voice source with a probability value determined as a function of the comparison result. In this way, the phase delay of the at least one phase delay element can be set such that at the location of the voice source a receiving maximum is obtained for the superposition signal. Consequently, voice signals from this voice source are received with a high sensitivity, whereas disturbance signals from other sound sources are excluded.

A special advantage comprises that the signal processing unit sets the phase delay of the at least one phase delay element as a function of the location of the identified voice source such that at the site of the voice source a receiving maximum results for the superposition signal. The setting of the phase delay of the at least one phase delay element in this way takes place automatically without intervention by a user, wherein the location of the greatest sensitivity, additionally, can be adaptively

tracking the location of the voice source. This represents a considerable improvement of the operating comfort for the user.

Drawing

An embodiment example of the invention is depicted in the drawing and will be explained in further detail in the following description. In the drawing show:

Fig. 1 a block circuit diagram of an electric apparatus, with the capacity for voice input, with two microphones, whose electric output signals are superimposed without phase shift, and

Fig. 2 a block circuit diagram of an electric apparatus with at least three microphones whose electric output signals can be superimposed with a phase shift.

Description of the Embodiment Example

In Figure 1 an electric apparatus with the capacity for voice input developed as a telecommunication end apparatus is characterized by 1. The telecommunication end apparatus 1 comprises an addition element 20 and a voice processing unit 70. An output 97 of the addition element 20 is connected to an input 107 of the voice processing unit 70. An output 108 of the voice processing unit 70 is connected to a telecommunication network not shown in Figure 1. To the telecommunication end apparatus 1 is connected a first microphone 5 and a second microphone 10. An output 104 of the first microphone 5 is connected to a first noninverting input 87 of the addition element 20 and an output 105 of the second microphone 10 is connected to a second noninverting input 90 of the addition element 20. At the same distance from the two microphones 5, 10 according to Figure 1 a sound source 55 is disposed developed as a loudspeaker, which outputs voice signals. The voice signals are

received by the two microphones 5, 10 according to the arrows drawn in dashed lines in Figure 1. The sound source 55 developed as a voice source can be, for example, the organ of speech of a user of the telecommunication end apparatus 1. The microphones 5, 10 transduce the received voice signals into electric signals and conduct them to the addition element 20, where they are superimposed through simple addition. Since the voice source 55 is equidistant from the two microphones 5, 10, the voice signal output by it, is weighted doubly in the addition element 20 due to the superposition of the electric output signals of the microphones 5, 10. The voice source 55 consequently is located at a location for which the superposition signal at output 97 of the addition element 20 yields a sensitivity or receiving maximum. The locations of the receiving maxima repeat themselves at the distance of the wavelength of the signal. Since voice represents a statistically distributed frequency mixture, on average only one receiving maximum occurs in the geometric center between the two microphones 5, 10 corresponding to a dashed line 200 in Figure 1.

In Figure 2 is represented a voice input-capable electric apparatus 1 according to the invention developed as a telecommunication end apparatus. It comprises a first phase delay element 30, a second phase delay element 35, a third phase delay element 40 and a fourth phase delay element 45. The telecommunication end apparatus 1 comprises moreover a signal processing unit 50, a voice analysis device 60 and a storage unit 65. In the telecommunication end apparatus 1, further, a first addition element 20 and a second addition element 25 as well as a voice processing unit 70 is provided. To the telecommunication end apparatus 1 is connected a first microphone 5, a second microphone 10 and a third microphone 15. One output 104 of the first microphone 5 is connected to a first input 85 of the first phase delay element 30 and to a first input 75 of the signal processing unit 50. An output 86 of the first phase delay element 30 is connected to a first noninverting input 87 of the first addition element 20. An output 105 of the second microphone 10 is connected to a first input 88 of the second phase delay element 35 and to a second input 76 of the signal processing unit 50. An output 89 of the second phase delay element 35 is connected to a second noninverting input 90 of the first addition element 20. An

output 106 of the third microphone 15 is connected to a first input 94 of the fourth phase delay element 45 and to a third input 77 of the signal processing unit 50. An output 95 of the fourth phase delay element 45 is connected to a first noninverting input 96 of the second addition element 25. A further microphone, not depicted in Figure 2, can be connected with its output, via a connection line shown in Figure 2 in dashed lines, to a first input 91 of the third phase delay element 40 and to a fourth input 78 of the signal processing unit 50. An output 92 of the fourth phase delay element 40 is connected to a second noninverting input 93 of the second addition element 25. An output 97 of the first addition element 20 is connected to a third noninverting input 98 of the second addition element 25. An output 99 of the second addition element 25 is connected to an input 107 of the voice processing unit 70. An output 108 of the voice processing unit 70 is connected to a telecommunication network, not shown in Figure 2. The voice processing unit 70 according to Figure 1 and Figure 2 has the task of preparing the superimposed electric voice signals for the transmission in the telecommunication network and to output them to it. If necessary, further microphones can be connected to the telecommunication end apparatus 1 and via corresponding phase delay and addition elements be superimposed with the remaining electric voice signals and supplied to the voice processing unit 70. The output signals of these further microphones are also to be supplied additionally to the signal processing unit 50. In the embodiment example depicted in Figure 2, which is provided for the connection of maximally four microphones, the signal processing unit 50 comprises a fifth input 79, which is connected to an output 110 of the storage unit 65. To the signal processing unit 50, further, a voice analysis device 60 is connected for mutual data exchange. Furthermore, a first output 81 of the signal processing unit 50 is connected to a second input 100 of the first phase delay element 30. A second output 82 of the signal processing unit 50 is connected to a second input 101 of the second phase delay element 35. A third output 83 of the signal processing unit 50 is connected to a second input 102 of the third phase delay element 40. A fourth output 84 of the signal processing unit 50 is connected to a second input 103 of the fourth phase delay

element 45. Furthermore, in Figure 2, again a sound source 55 is shown developed as a loudspeaker, which outputs voice signals and which can represent, for example, the organ of speech of a user. The three microphones 5, 10, 15 receive according to the dashed arrows in Figure 2 the voice signals of the sound source 55 developed as a voice source. According to Figure 2, the voice source 55 is located at a geometric location represented by a dashed line 200, which, in contrast to the configuration according to Figure 1, no longer forms the geometric center between the three microphones 5, 10, 15 such that the three microphones 5, 10, 15 are at a different distance from the voice source 55.

If a noncentral directional effect is to be attained, the location for which the superposition signal of the electric voice signals yields a receiving maximum, can be preset through the suitable selection of the phase delay of the individual phase delay elements 30, 35, 40, 45. Thereby, a receiving maximum can be attained even for the noncentral disposition of the voice source 55 according to Figure 2. Depending on the location of the voice source 55, it may already be sufficient to delay only a single microphone output signal in its phase such that, limited to this application case, only one phase delay element would be required. By using a phase delay element for each microphone, however, greater flexibility is given for the presetting of the location for the voice source 55, in which the superposition signal at input 107 of the voice processing unit 70 yields a receiving maximum. Since, for setting a receiving maximum, the attachment locations of the microphone 5, 10, 15 are critical, in a preset location for the voice source 55 a receiving maximum can be set for the superposition signal through suitable disposition of the microphones 5, 10, 15 as well as through suitable selection of the phase delays of the phase delay elements 30, 35, 40, 45. However, if the attachment locations of the microphones 5, 10, 15 are not changeable, the receiving maximum for the superposition signal can only be attained by variation of the phase delays of the phase delay elements 30, 35, 40, 45.

Through suitable selection of the attachment locations for microphones 5, 10, 15 and the phase delays of the phase delay elements 30, 35, 45 connected with microphones 5, 10, 15, the receiving sensitivity of the telecommunication end

apparatus 1 for specific regions can be increased or decreased, such that disturbing sound sources in the region of lesser sensitivity are substantially excluded and useful sound sources can be received with improvement in the region of increased sensitivity. For each sound source, the receiving sensitivity can therein be defined within a preset region.

From the output signals supplied to it of the microphones 5, 10, 15 the signal processing unit 50 can optionally also calculate a three-dimensional image of the sound environment such that all sound sources can be determined locally. When using only two microphones, only a two-dimensional image of the sound environment can be determined. When using more than three microphones, the accuracy in the location determination of the sound sources can be increased, however, greater arithmetic expenditures are required. By means of the voice analysis device 60, a comparison of parameters of the electric microphone output signals with the voice parameters deposited in the storage unit 65 is possible. The signal processing unit 50 determines therein, as a function of the comparison result for each sound source detected in the sound environment, a value which indicates the probability with which the particular sound source has been recognized as a voice source. The voice source with the highest probability value is subsequently identified as the voice source. With this information, the phase delays can be set with the phase delay elements 30, 35, 45 connected with microphones 5, 10, 15 such that at the location of the sound source identified as voice source a receiving maximum is obtained for the superposition signal. The other sound sources are thus substantially excluded as disturbance sources. The corresponding setting of the phase delays can also take place automatically through the signal processing unit 50, such that an adaptation of the phase shifts of the phase delay elements 30, 35, 45 connected with microphones 5, 10, 15 to a changing locations of the sound source identified as voice source is possible such that, in spite of a relative movement between the voice source 55 and the telecommunication end apparatus 1 or the microphones 5, 10, 15, at the location of the voice source 55 for the superposition signal a receiving maximum is maintained.

In the case of several sound sources recognized with high probability values, the user can also preset a sound source as voice source. This is of advantage, for example, if the telecommunication end apparatus 1 is integrated into a car radio and the driver as well as also a passenger has to be considered as the voice source. Through appropriate changes of the phase delays of the phase delay elements 30, 35, 45 connected with microphones 5, 10, 15, the driver or a passenger can be selected as voice source 55 so that for the location of the selected voice source a receiving maximum is set for the superposition signal.

If the microphones 5, 10, 15 are components of a hand set-free speaking device of the telecommunication end apparatus 1, it is possible to receive voice signals of a voice source 55 directed location specific and nearly without disturbances. The voice comprehensibility of the voice signals received by the hand set-free speaking device is thereby considerably improved.

The invention is not limited to a telecommunication end apparatus 1, but rather applicable to all electric apparatus with the capacity for voice input. This can be, for example, also apparatus which comprise a voice control. In this case, the voice processing unit 70 serves for evaluating and initiating voice commands. Since in the evaluation of voice commands the disturbance free reception is critical, the separation of useful and noise signals according to the invention makes possible maximum error-free detection of the voice commands without special mechanical auxiliary means, such as for example directional microphones or special filter algorithms, for the elimination of noise signals being required.

If the electric apparatus 1 with the capacity for voice input is implemented as a telecommunication end apparatus, due to the adaptive tracking of the receiving maximum for the superposition signal, with a relative movement between the telecommunication end apparatus 1 and the voice source 55 it is not required that the telecommunication end apparatus 1 is disposed stationarily. Therefore, the invention is also applicable to radio equipment, mobile telephones, wireless telephones and the like. The same applies to mobile electric apparatus with the capacity for voice input with voice control. Electric apparatus capable of voice input with voice control can

be, for example, car radios, personal computers, and the like, however, also wire or wireless telecommunication end apparatus.

It is also possible to realize the phase delays and additions, instead of with discrete assemblies, in the signal processing unit 50 or a separate signal processing unit. As signal processing unit can be used, for example, a digital signal processor.

The process according to the invention and the electric apparatus according to the invention can be used generally for optimizing the reception of any acoustic signal such that restriction to voice-input capable electric apparatus is not required. A speech analysis is also not required in this case. For the selection of a sound source as useful sound source, correspondingly suitable criteria must in this case be selected, which must correspondingly be taken into account by the signal processing unit 50. It is also possible to provide allowing a user to select at an input unit a sound source as useful sound source. The sound sources not selected as useful sound source are subsequently excluded by means of suitable phase delays. Therein the phase delays are set by the signal processing unit 50 such that an adaptive tracking of the receiving sensitivity takes place as a function of the location of the useful sound source, wherein the disturbance sound sources are adaptively excluded as a function of the location.

Patent Claims

1. Process for optimizing the reception of acoustic signals, **characterized in that** signals received by at least two microphones (5, 10, 15) are transduced into electric output signals, that signals derived from the electric output signals are superimposed, wherein at least one electric signals, before the superposition, is delayed in its phase and that the at least one phase delay is selected such that, as a function of the location of the connected microphones (5, 10, 15) for at least one preset site (200) the amplitude of the superposition signal is within a preset range if at the at least one preset site (200) a sound source (55) outputs acoustic signals.

2. Process as claimed in claim 1, characterized in that, as a function of the amplitude of the signals derived from the electric output signals, location co-ordinates of at least one sound source (55) are determined.

3. Process as claimed in claim 1 or 2, characterized in that from the signals derived from the electric output signals, voice parameters are derived, that the voice parameters are compared with preset voice parameters, and that a sound source (55) is identified as voice source with a probability value determined as a function of the comparison result.

4. Process as claimed in claim 1, 2, or 3, characterized in that the phase of the at least one electric signal before the superposition, is delayed as a function of the location of the identified sound source (55) such that at the site of the sound source (55) a receiving maximum results for the superposition signal.

5. Electric apparatus (1), in particular telecommunication end apparatus, characterized in that at least two microphones (5, 10, 15) are connectable to the electric apparatus (1), which transduce received acoustic signals into electric output signals, that at least one means (20, 25), in particular an addition element, is provided which superimposes the signals derived from the electric output signals of connected microphones (5, 19, 15), wherein at least one phase delay element (30, 35, 40, 45) is provided which delays an electric signal in its phase before the superposition, and that the at least one phase delay of the at least one phase delay element (30, 35, 40, 45) is selected such that, as a function of the location of the connected microphones (5, 10, 15) for at least one preset site (200) the amplitude of the superposition signal is within a preset range if at the at least one preset site (200) a sound source (55) outputs acoustic signals.

6. Electric apparatus (1) as claimed in claim 5, characterized in that different phase delays of the at least one phase delay element (30, 35, 40, 45) can be set.

7. Electric apparatus (1) as claimed in claim 5 or 6, characterized in that a signal processing unit (50) is provided to which are supplied electric signals of microphones (5, 10, 15) and which, as a function of the amplitudes of the electric signals, determines location coordinates of at least one sound source (55).

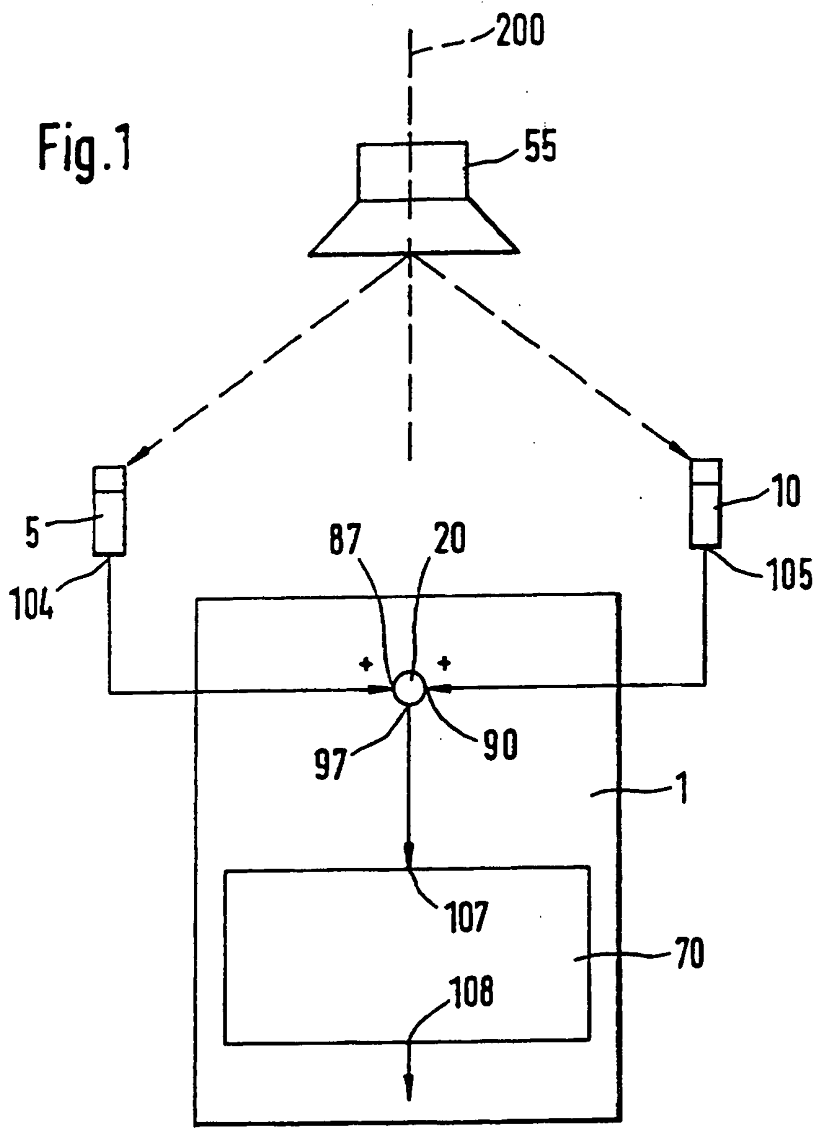
8. Electric apparatus (1) as claimed in claim 5, 6, or 7, characterized in that with the signal processing unit (50) is associated a voice analysis device (60) and that the voice analysis device (60) carries out a comparison of parameters of the electric signals with voice parameters stored in an associated storage unit (65) and with a probability value, determined as a function of the comparison result, identifies one sound source (55) as voice source.

9. Electric apparatus (1) as claimed in one of claims 5 to 8, characterized in that the signal processing unit (50) sets the phase delay of the at least one phase delay element (30, 35, 40, 45) as a function of the location of the identified sound source (55) such that at the site of the sound source (55) a receiving maximum results for the superposition signal.

2 sheets of drawings enclosed

Nummer: DE 197 41 596 A1
Int. Cl.⁶: H 04 R 1/20
Offenlegungstag: 25. März 1999

Fig.1



Nummer: DE 197 41 596 A1
Int. Cl.[6]: H 04 R 1/20
Offenlegungstag: 25. März 1999

Fig.2